

The Cognitive Features of Programming Language and Natural Language

Wen Xu, Fangqu Xu, Xiaohui Zou, Zhenlin Xu

► **To cite this version:**

Wen Xu, Fangqu Xu, Xiaohui Zou, Zhenlin Xu. The Cognitive Features of Programming Language and Natural Language. 2nd International Conference on Intelligence Science (ICIS), Nov 2018, Beijing, China. pp.184-190, 10.1007/978-3-030-01313-4_19 . hal-02118845

HAL Id: hal-02118845

<https://hal.inria.fr/hal-02118845>

Submitted on 3 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



The Cognitive Features of Programming Language and Natural Language

[0000-0002-5577-8245]

Wen Xu, Fangqu Xu, Xiaohui Zou, Zhenlin Xu

* Shanghai Shenyue Software Technology Co., Ltd.
Room B10, North D., Bld. 8, No. 619 Longchang Rd.,
Shanghai 200090,
China
cindyxu@geomking.com

** Shanghai Shenyue Software Technology Co., Ltd.
Room B10, North D., Bld. 8, No. 619 Longchang Rd.,
Shanghai 200090,
China
xufangqu@geomking.com

*** Sino-American Saerle Research Center
Tiangongyuan, Paolichuntianpai Building 2, Room 1235, Daxing
Beijing 102629
China
949309225@qq.com

**** Shanghai Shenyue Software Technology Co., Ltd.
Room B10, North D., Bld. 8, No. 619 Longchang Rd.,
Shanghai 200090,
China
hyperien@qq.com

ABSTRACT. The purpose of this paper is to establish the mechanism of mutual transformation between programming language and natural language through the cognitive features of programming language and natural language which enables computers to read natural language and furthermore, to think like a human brain to achieve the purpose of human-computer interaction. The method is: First, establish a data dictionary, a vocabulary table, a sentence table and a bilingual table so that natural language and programming language can establish a one-to-one, one-to-many conversion relationship. Second, according to the data dictionary and bilingual table, establish the thinking tree model, and implement the traversal of natural language (a particular mathematical problem) through programming language. Third, establish a selection mechanism so that,

through programming language, the computer knows how to make different choices and responses corresponding to different natural languages. The result is: Through the exploration of the cognitive features of programming language and natural language, establish the mechanism of the mutual transformation between programming language and natural language. The significance is: through the use of two different languages, it can establish the mechanism of human-computer interaction. For the one-to-one situations, it can be automatically converted based on rules. On the other hand, for the one-to-many situation, it can realize machine learning, moreover, deep learning based on statistics. Thus, it lays the foundation for the application of artificial intelligence in the field of education.

KEYWORDS: natural language; programming language; cognitive features; artificial intelligence; deep learning

1 Introduction

Artificial Intelligence is a new technical science for researching and developing theories, methods, technologies, and application systems for simulating, extending, and expanding human intelligence. Artificial intelligence is a branch of computer science. It attempts to understand the essence of intelligence and produces a new intelligent machine that can respond in a similar manner to human intelligence. Research in this area includes robots, language recognition, image recognition, Natural language processing, expert systems, etc.

The most essential attribute of human intelligence is thinking. This is because all human achievements in the process of social progress and inheritance, especially innovation and creative achievements, are all thought out and the results of thinking. Since the basis of thinking is language, and the process of thinking is also expressed through language, the study of human intelligence and the study of artificial intelligence will certainly bring language research a basic and key research role. It is not only an entry point, but also a breakthrough point. .

The purpose of this paper is to establish the mechanism of mutual transformation between programming language and natural language through their cognitive features, so that the computer can read natural language and further think like the human brain, thus achieving the purpose of human-computer interaction. .

2 Cognitive features of programming and natural language

Intelligent machines that can respond in a similar manner to human intelligence must have their own language system. The source of this language system is natural language and programming language.

Natural language is the form that human uses to express the information that needs to be expressed in sound and words (the mother tongue of the country and the nation). Natural language is the most commonly used, popular, and important form of human expressions for thinking activity, thinking process, and thinking achievement. It is

also the most basic and most important form of works publishing by scientists, experts, and scholars (Yixin Zhong, 2004).

Taking the application of the “Thinking – student’s thinking process visualization and evaluation system (Abbr. Thinking)” as an example, for any test question, in order to achieve the goal of recording, display and evaluation of every single step by student’s thinking process in computer, it’s necessary that all the thinking nodes of this problem should be written out step by step, the situation that every thinking node student may think of should be listed in full or as much as possible. The result could build such a “Thinking Tree” map. Obviously at this time only natural language could be used (of course, including graphic language, symbolic language, etc.).

Example (Fangqu Xu, Wen Xu, 2017):

In picture, parabola $y=ax^2-8ax+12a$ ($a<0$) intersects with X axis at points A, B (point A is at left side of point B). Another point C in parabola is in the first quadrant, so that $\angle ACB$ is a right angle, and $\triangle OCA \sim \triangle OBC$.

1. Find the length of line segment OC
2. Find the parabolic function
3. Judge if there is point P in X axis, so that $\triangle BCP$ is an isosceles triangle? If it exists, please list all possible coordinates of P point. If it doesn’t exist, please give the reason.

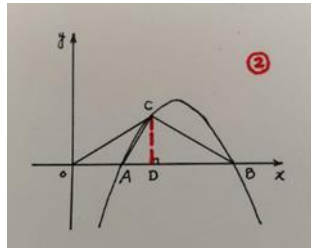


Fig. 1. Example figure

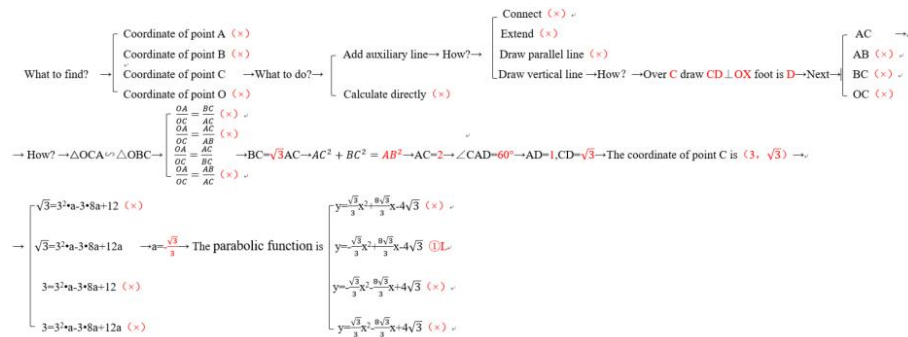
4. Only the solution idea needs to be answered

Where to start thinking? →

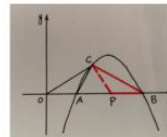
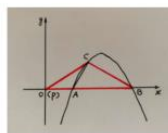
- Find coordinates of point A and B → How? → $ax^2-8ax+12a=0 \rightarrow x_1=2, x_2=6 \rightarrow$ Coordinate of point A is (2,0), coordinate of point B is (6,0) →
- Find coordinate of point C (×) ✓
- Find value of a (×) ✓
- Find the vertex coordinates of the parabola (×) ✓

→ $\begin{cases} OA=2 \\ OB=6 \end{cases}$ →What should be used for following proof? → $\triangle OCA \sim \triangle OBC \rightarrow OC^2 = OA \cdot OB \rightarrow OC=2\sqrt{3}$.

5. Only the solution idea needs to be answered



6. Only the solution idea needs to be answered



Obviously, in the above example, the natural language can be used clearly, accurately, and step by step to express the thinking process. However, such language can not make the computer “understand”, and can not be used to display the contents in computer directly.

In order to let computer “understand” natural language, a language system should be established by computer so that natural language could be run and displayed. That is the programming language (Tucker, Noonan, 2009).

Computers need to deal with and demonstrate a natural language system, and run a programming language system. Therefore, a common working mode of natural language and programming language on a computer at the same time is formed.

Running both language systems in computer accurately, harmoniously and efficiently, achieving the goal of recording, displaying and evaluating the thinking process is obviously a challenge to human wisdom too.

Since natural language and programming language are used together on a computer, the first step is to establish the correspondence between natural language and programming language. The establishment and formation of this correspondence involves two aspects of the roadmap (Daniel Jurafsky, James H. Martin, 2005):

- In first way, establish a one-to-one correspondence, and then promote it to a one-to-many correspondence. For the already established one-to-one correspondence between natural language and programming language, it’s possible to realize the automatic conversion by computer. And then on this basis the one-to-many

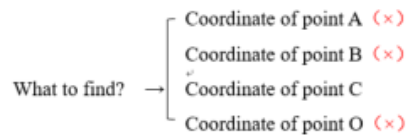
correspondence between natural and programming language could be conversed automatically by computer.

- In second way, it starts from the "basic element" of natural language and programming language, constructs vocabularies and sentence tables, and bilingual tables based on established correspondences to form a data dictionary.

In "Thinking" software, the basic element is not only the word and sentence, but also the basic element of the thinking process. For example: by the thinking process of congruent triangles, there is a fixed thinking mode: in two triangles, if we can get three conditions that could lead to the congruent triangles judgement, we can prove that these two triangles are congruent triangles, and furthermore their corresponding edges or corresponding angles are proven to be equal. The whole part mentioned about congruent triangles builds then a basic element of thinking.

During the thinking of a problem, goes a wrong way or trial and error could often happen, and this also needs to be recorded, displayed and evaluated. So on the thinking node, it is necessary to design students' possible wrong thinking activities. Since the common mistakes that students could make in a certain thinking node are basically also determined, the choices designed on such thinking nodes can also become a basic element of thinking.

For example, in the above mentioned example we extract one choice option:



Per natural language, it could be described like following:

What should be found in this case? There are four options: coordinate of point A, coordinate of point B, coordinate of point C, coordinate of point O, the correct answer is coordinate of point C.

Per programming language (take JavaScript as example), it could be described like following:

```
{
  code: 's010',
  title: 'What item should be found here',
  sType: 'S',
  reviewPoint: "",
  choices: [
    { match:'A', title: 'Coordinate of point A'},
    { match:'B', title: 'Coordinate of point B'},
    { match:'C', title: 'Coordinate of point C'},
    { match:'D', title: 'Coordinate of point D'},
  ],
  branches: [
    { match: 'C', correct: true, next: 's030' },
  ]
},
```

These two different languages, natural language and programming language, express the same meaning under different forms in different scenarios. Therefore, we can establish a correspondence between them.

After the data dictionary has been established and obtained, the second step is to establish the “thinking tree” model based on the data dictionary, so that every mathematical problem could be gone through by programming language. .

The ability to evaluate the thinking process on computers is because we have built a “thinking tree” model. The thinking tree for every concrete mathematical problem starts with asking such questions: what nature does the thinking start from? What is the way to think about the question? What should the first step of the question be to seek (do)? Etc. This is the first thought node. In this thinking node, we can carry out a diffusion thinking, list all the possible situations that you think out, and then go into the second thinking node along each path, and the thinking process can be developed in turn until the problem is solved. In this process, a thinking tree is formed.

Obviously, such a thinking tree is expressed by natural language.

The next step is to use programming language to go through all natural language used in the thinking tree. It's necessary to understand the natural language used in the thinking tree, especially the basic elements which formulate the thinking tree.

The third step is to establish a selection mechanism through programming and programming languages. The computer can “understand” the natural language used in the thinking tree through the correspondence between the programming language and natural language, and compare the various information expressed in the natural language and make a judgment, reaction and choice. In the “Thinking” software, the thinking tree is invisible to the teachers and the tested students, because the thinking tree has been compiled into the computer program through the programming language, but the computer can faithfully record every step of thinking of the tested student.

Which step he thought of, thought wrong, what's the reason? What's the reason if he went around in circles and went back? As long as the students who have been tested have a thinking behaviour, there are records on the computer, and these records can be evaluated. Thus, the students' thinking process, thinking ability, thinking level and thinking quality can be evaluated.

3 Conclusion

The result of this paper is to establish the mechanism of mutual transformation between programming language and natural language through the research and exploration of the cognitive features of the programming language and natural language, and finally make it possible to evaluate the thinking ability based on the thinking process.

The significance of this paper is reflected in the promotion of educational development and reform. It will change the fundamental problem of traditional education for a long time, actually the evaluation of the education quality bases on the correctness of test answers solely. As long as the answer is correct, there is full score, whether the answer is thought out or remembered, even plagiarized. In the application of the "Thinking", even if the students could do every exercise in one test paper, let them test in the process of thinking, thought and not thought, think out and do not

think out, one evaluation is clear, thus the evaluation could be developed from the evaluation of answer solely to the evaluation of process.

The basic thinking element, the thinking tree model and the data dictionary established in this paper can promote the computer to automatically generate new exercises bank and new thinking tree bank, and spread it into other subject areas so that a more complete automatic teaching system, automatic proposition system, automatic scoring system could be developed, thus the real education artificial intelligence could be realized.

By using two different languages together, a human-machine interaction mechanism can be established. Among them, one-on-one situations can be automatically converted based on rules, and one-to-many situations can be converted based on machine learning or even deep learning, so that a solid base for education artificial intelligence could be established. ..

4 References

1. Yixin Zhong, "Comprehensive Information Based Methodology for Natural Language Understanding", Journal of Beijing University of Posts and Telecommunications, vol. 27, No.4, 2004
2. Xiaolong Wang, "Natural language processing by computer", Tsinghua University Press, 2005
3. Daniel Jurafsky, James H. Martin, "Speech and Language Processing", Publishing House of Electronics Industry, 2005
4. Allen B. Tucker, Robert E. Noonan, "Programming Languages Principles and Paradigms, Second Edition", Tsinghua University Press, 2009
5. Fangqu Xu, Wen Xu, "Transparent Geometry - New Practice of Internet + Planar Geometry", Shanghai Education Publishing House, 2017